



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/724,170	12/01/2003	Maurcen Caudill	000479.00114	9385
22907	7590	12/19/2006	EXAMINER	
BANNER & WITCOFF 1001 G STREET N W SUITE 1100 WASHINGTON, DC 20001			MYINT, DENNIS Y	
			ART UNIT	PAPER NUMBER
			2162	
SHORTENED STATUTORY PERIOD OF RESPONSE		MAIL DATE	DELIVERY MODE	
3 MONTHS		12/19/2006	PAPER	

Please find below and/or attached an Office communication concerning this application or proceeding.

If NO period for reply is specified above, the maximum statutory period will apply and will expire 6 MONTHS from the mailing date of this communication.

Office Action Summary	Application No.	Applicant(s)
	10/724,170	CAUDILL ET AL.
	Examiner Dennis Myint	Art Unit 2162

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on 10 October 2006.
- 2a) This action is FINAL. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 81-98 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) Claim(s) _____ is/are allowed.
- 6) Claim(s) 81-98 is/are rejected.
- 7) Claim(s) _____ is/are objected to.
- 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on 01 December 2003 is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
a) All b) Some * c) None of:
1. Certified copies of the priority documents have been received.
2. Certified copies of the priority documents have been received in Application No. _____.
3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)	4) <input type="checkbox"/> Interview Summary (PTO-413)
2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)	Paper No(s)/Mail Date. _____
3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____	5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)
	6) <input type="checkbox"/> Other: _____

DETAILED ACTION

1. This communication is responsive to Applicant's Amendment, filed on 10 October 2006.
2. In the Amendment filed on 10 October 2006, claims 54-58 and 70-80 were cancelled. Claims 81-98 are newly added. Claims 81, 87, and 93 are independent claims. As such, claims 81-98 are pending in this case. This office action is made final.

Response to Arguments

3. The applicant's arguments filed on 05 September 2006 have been fully considered but are not persuasive.

On Page 7 of Applicant's arguments, Applicants argued that *the combination of call and Maarek fails to teach or suggest each and every feature of Applicants' claim 81. Applicants' new claim 81 recites, among others, "identifying at least one predicate and argument in said set of document predicate structures".*

In response, it is pointed out that Maarek in view of Call teaches the limitations of Claim 81 as "identifying at least one predicate and argument in said set of document predicate structures (Maarek , Page 5, Paragraph 4)" and "estimating conceptual nearness of two of said document predicate structures in said set of predicate structures" (Maarek, Page 7 Last Paragraph and Page First Paragraph, i.e. *Instead of the typical use of single words as indexing units, our indexing unit consists of a pair of words that are linked by a lexical affinity (LA)*). Maarek teaches a method for clustering

using profile (word vectors) and document vectors wherein, “predicate structures” (Maarek et al., Page 5, Paragraph 4, i.e. *An indexing term can be a single term (possibly represented by a canonical form such as its morphological root, lemma or stem), or it may take more complex form such as phrases, syntactic constructs or lexical constructs.*) “are vectorized” (Maarek et al., Page 5, Paragraph 4, i.e., *For example, if the indexing units are single words, then each word represents an axis in a high-dimensional vector space, where the dimension is equal to the number of words in the collection.* and Page 7 Last Paragraph and Page First Paragraph, i.e. *Instead of the typical use of single words as indexing units, our indexing unit consists of a pair of words that are linked by a lexical affinity (LA)*). Call teaches the limitations: “set of document predicate structures” (Call, Paragraph 0117, i.e. *Items may be organized into set which consist simply of an ordered collection of item numbers which are gathered in accordance with some criteriaand The time in a set need not be of the same type, however, but may be collected in a single set based on the fact that they share some common attribute. Thus, times of type “apple” and of type “orange” may be collected together to form a set named “treefruit”.*), “by a predicate key that is an integer representation” (Call, Paragraph 0070, i.e., *a general purpose database program which stores natural language text and a rich variety of other typed data in an array of integers subdivided into data elements called items..*)) and “by subtracting corresponding ones of said predicate keys” (Note that, in the method of Call, data elements/concepts are identified by integers and data elements/concepts of similar attribute are organized into

Art Unit: 2162

sets. Thus, it is inherent in Call's method that conceptual nearness is obtained by subtracting corresponding one of predicate keys).

Duplicate Claims Warning

4. Applicant is advised that should claim 81-86 be found allowable, claims 87-92 will be objected to under 37 CFR 1.75 as being a substantial duplicate thereof (*Claim 87 is a duplicate of claim 81; Claim 88 is a duplicate of Claim 82; Claim 89 is a duplicate of Claim 83; Claim 90 is a duplicate of Claim 84; Claim 91 is a duplicate of Claim 85; and Claim 92 is a duplicate of Claim 86*). When two claims in an application are duplicates or else are so close in content that they both cover the same thing, despite a slight difference in wording, it is proper after allowing one claim to object to the other as being a substantial duplicate of the allowed claim. See MPEP § 706.03(k).

Claim Rejections - 35 USC § 101

5. 35 U.S.C. 101 reads as follows:

Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

6. Claims 81-98 are rejected under 35 U.S.C. § 101 because the claimed invention is directed to non-statutory subject matter.

Claims 81, 87, and 93 are not statutory because the claims do not require any physical transformation and the invention as claimed does not produce a useful and

Art Unit: 2162

tangible result in view of MPEP 2106 (IV)(C)(2)((B))((2))(a) and (b)¹.

Claims 82-84 are rejected under 35 U.S.C. § 101 because of their dependency on claim 81.

Claims 88-90 are rejected under 35 U.S.C. § 101 because of their dependency on claim 87.

Claims 94-98 are rejected under 35 U.S.C. § 101 because of their dependency on claim 93.

Claim Rejections - 35 USC § 103

7. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

8. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein

¹ MPEP 2106 (IV)(C)(2)((B))((2))(a) and (b):

For an invention to be "useful" it must satisfy the utility requirement of section 101. The USPTO's official interpretation of the utility requirement provides that the utility of an invention has to be (i) specific, (ii) substantial and (iii) credible. MPEP § 2107 and Fisher, 421 F.3d at 1372, 76 USPQ2d at 1230 (citing the Utility Guidelines with approval for interpretation of "specific" and "substantial"). In addition, when the examiner has reason to believe that the claim is not for a practical application that produces a useful result, the claim should be rejected, thus requiring the applicant to distinguish the claim from the three 35 U.S.C. 101 judicial exceptions to patentable subject matter by specifically reciting in the claim the practical application. In such cases, statements in the specification describing a practical application may not be sufficient to satisfy the requirements for section 101 with respect to the claimed invention. Likewise, a claim that can be read so broadly as to include statutory and nonstatutory subject matter must be amended to limit the claim to a practical application. In other words, if the specification discloses a practical application of a section 101 judicial exception, but the claim is broader than the disclosure such that it does not require a practical application, then the claim must be rejected.

The tangible requirement does not necessarily mean that a claim must either be tied to a particular machine or apparatus or must operate to change articles or materials to a different state or thing. However, the tangible requirement does require that the claim

were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

9. Claims 81-83, 86, 87-89, and 92 are rejected under 35 U.S.C. 103(a) as being unpatentable over of Maarek et al. (hereinafter "Maarek") ("Ephemeral Document Clustering for Web Applications", IBM Research Report, RJ 10186, April 2000) in view of Call (hereinafter "Call") (U.S. Patent Application Publication Number 2002/0143521).

As per claim 81, Maarek is directed to a method for vectorizing a set of document predicate structures (Maarek et al., Page 5, Paragraph 4) and teaches the limitations:

"identifying at least one predicate and argument in said set of document predicate structures (Maarek , Page 5, Paragraph 4)" and "estimating conceptual nearness of two of said document predicate structures in said set of predicate structures" (Maarek, Page 7 Last Paragraph and Page First Paragraph, i.e. *Instead of the typical use of single words as indexing units, our indexing unit consists of a pair of words that are linked by a lexical affinity (LA)*). Maarek teaches a method for clustering using profile (word vectors) and document vectors wherein, "predicate structures" (Maarek et al., Page 5, Paragraph 4, i.e. *An indexing term can be a single term (possibly represented by a canonical form such as its morphological root, lemma or*

stem), or it may take more complex form such as phrases, syntactic constructs or lexical constructs.) “are vectorized” (Maarek et al., Page 5, Paragraph 4, i.e., For example, if the indexing units are single words, then each word represents an axis in a high-dimensional vector space, where the dimension is equal to the number of words in the collection. and Page 7 Last Paragraph and Page First Paragraph, i.e. *Instead of the typical use of single words as indexing units, our indexing unit consists of a pair of words that are linked by a lexical affinity (LA)).*

Maarek does not explicitly teach the limitations: “set of document predicate structures”, “by a predicate key that is an integer representation” and “by subtracting corresponding ones of said predicate keys”.

Call teaches the limitations:

“set of document predicate structures” (Call, Paragraph 0117, i.e. *Items may be organized into set which consist simply of an ordered collection of item numbers which are gathered in accordance with some criteriaand The time in a set need not be of the same type, however, but may be collected in a single set based on the fact that they share some common attribute. Thus, times of type “apple” and of type “orange” may be collected together to form a set named “treefruit”).*

“by a predicate key that is an integer representation” (Call, Paragraph 0070, i.e., *a general purpose database program which stores natural language text and a rich variety of other typed data in an array of integers subdivided into data elements called items..)) and*

"by subtracting corresponding ones of said predicate keys" (Note that, in the method of Call, data elements/concepts are identified by integers and data elements/concepts of similar attribute are organized into sets. Thus, it is inherent in Call's method that conceptual nearness is obtained by subtracting corresponding one of predicate keys).

At the time the invention was made, it would have been obvious to a person of ordinary skill in the art to add the feature of Call for storing data elements as sequences of integers with the method of Maarek for vectorizing phrases/pairs of concepts (words) so that the combined method would be a method for "vectorizing a set of document predicate structures, which comprises the steps of identifying at least one predicate and argument in said set of document predicate structures by a predicate key that is an integer representation, and estimating conceptual nearness of two of said document predicate structures in said set of document predicate structures by subtracting corresponding one of said predicate keys." One would have been motivated to do so because computing machines could more efficiently manipulate *characters/symbols* when they are *represented by integers, Booleans, floating points, logical values, or the like* (Call, Paragraph 0009).

Referring to claim 82, Maarek in view of Call as applied to claim 81 above teaches the limitation:

"the method further comprising the further step of constructing multi-dimensional vectors using said integer representation" (Maarek et al., Page 5, Paragraph 4, i.e., *For*

example, if the indexing units are single words, then each word represents an axis in a high-dimensional vector space, where the dimension is equal to the number of words in the collection and Page 7 Last Paragraph and Page First Paragraph, i.e. *Instead of the typical use of single words as indexing units, our indexing unit consists of a pair of words that are linked by a lexical affinity (LA).*). Note that Maarek employs pairs of words and word vectors. Therefore, vectoring method of Maarek is multidimensional.

Referring to claim 83, Maarek in view of Call as applied to claim 81 above teaches the limitation:

"the method further comprising the further step of normalizing said multi-dimensional vectors" (Maarek et al., Page 5 Third Paragraph, i.e. *after normalization.*)

As per claim 86, Maarek teaches the limitation:

"wherein said set of document predicate structures are representations of logical relationships between words in a sentence" (Maarek, Page 7 Last Paragraph and Page First Paragraph, i.e. *Instead of the typical use of single words as indexing units, our indexing unit consists of a pair of words that are linked by a lexical affinity (LA)).*

Claim 92 is rejected on the same basis as claim 86.

Claims 87-89 are rejected on the same basis as claims 81-83 respectively.

10. Claim 84-85 and 90-91 rejected under 35 U.S.C. 103(a) as being unpatentable over Maarek in view of Call and further in view of Liddy et al. (hereinafter "Liddy") (U.S. Patent Number 5873056).

As per claim 84, Maarek in view of Call as applied to claim 81 above does not explicitly disclose the limitation: "comprising further step of identifying at least one query predicate structure by a second predicate key that is a second integer representation, and constructing second multi-dimensional vectors, for said at least one query predicate structure, using said integer representation".

Liddy teaches the limitation:

"comprising further step of identifying at least one query predicate structure by a second predicate key that is a second integer representation, and constructing second multi-dimensional vectors, for said at least one query predicate structure, using said integer representation" (Liddy et al., Column 5 Line 23-42). Liddy teaches a method for natural language processing with semantic vector representation, wherein queries are transformed into vectors to match document vectors (Liddy et al., Column 5 Line 23-42).

At the time the invention was made it would have been obvious to a person of ordinary skill in the art to add the feature of transforming queries into vectors as taught by Liddy et al. to the method of Maarek in view of Call so that, in the resultant method, query predicate structures will be included and the method would "further comprise the step of identifying at least one query predicate structure by a second predicate key that is a second integer representation, and constructing second multi-dimensional vectors, for said at least one query predicate structure, using said second integer

representation." One would have been motivated to do so in order to *enable retrieving of documents relevant to a query by matching a vector representing the query to the vectors representing documents*" (Liddy Column 1 Line 22-28).

Claim 85 is rejected on the same basis as claim 84.

Claims 90-91 are rejected on the same basis as claim 84-85 respectively.

11. Claim 93 and 95-98 are rejected under 35 U.S.C. 103(a) as being unpatentable over Maarek in view of Call and further in view of Liddy and further in view of Dorocak.

Referring to claim 93, Maarek in view of Call and further in view of Liddy is directed to "the method of constructing multi-dimensional vector representations for each document of a set of documents" (Maarek) and teaches the limitations:

"determining each predicate structure of one or more predicate structures M in each document of the set of documents" (Liddy, Column 5 Line 31-33 *The lexical database which is used determines the SFC's*), "said M predicate structures include a predicate and at least one argument" (Maarek, Page 5, Paragraph 4, i.e. *An indexing term can be a single term (possibly represented by a canonical form such as its morphological root, lemma or stem), or it may take more complex form such as phrases, syntactic constructs or lexical constructs.*); and

"identifying the predicate and the at least one argument in each of said M predicate structures by a predicate key that is an integer representation" (Call, Paragraph 0070, i.e., *a general purpose database program which stores natural*

language text and a rich variety of other typed data in an array of integers subdivided into data elements called items..);

However, the method of Maarek in view of Call and further in view of Liddy does not explicitly disclose the limitation “determining the fixed number of arguments q for vector construction”.

On the other hand, Dorocak teaches the limitation:

“determining the fixed number of arguments q for vector construction” (Page 104 Line 38-45). Dorocak teaches a method the specification of context-sensitive properties for programming languages, *wherein the number of arguments are less than the specified number of arguments, the unfilled argument positions are filled with zeroes* (Dorocak, Page 104 Line 38-45, i.e. *Where the number of arguments is less than the number of parameters specified in the corresponding definition, the argument list will be assumed to be filled out by arguments whose value is zero.*).).

At the time the invention was made, it would have been obvious to a person of ordinary skill in the art to add the feature of filling the unfilled arguments with zeros as taught by Dorocak to the method of Maarek in view of Call and further in view of Liddy et al. so that the resultant method, “would further comprise

determining a fixed number of arguments q for vector construction” (Dorocak, Page 104, *specified in corresponding definition*);

“constructing an N-dimensional vector representation of each document based upon the predicate and q arguments” (Maarek et al., Page 5, Paragraph 4, i.e., *For example, if the indexing units are single words, then each word represents an axis in a*

high-dimensional vector space, where the dimension is equal to the number of words in the collection. and Page 7 Last Paragraph and Page First Paragraph, i.e. *Instead of the typical use of single words as indexing units, our indexing unit consists of a pair of words that are linked by a lexical affinity (LA).).*

"wherein any predicate structure of said M predicate structures that includes less than q arguments fills unfilled argument positions with a numerical zero" (Dorocak, Page 104 Line 38-45, i.e. *Where the number of arguments is less than the number of parameters specified in the corresponding definition, the argument list will be assumed to be filled out by arguments whose value is zero.*).

One would have been motivated to do so in order to *provide a syntactic specification of the default attributes of a language by the description of modifications with are to be made the parsed form (syntactic tree) of instances of language.* (Dorocak, Page 101 Line 7-13).

As per claim 95, Maarek in view of Call and further in view of Liddy and further in view of Dorocak teaches the limitation:

"wherein conceptual nearness of two of said N-dimensional vector representation is estimated by subtracting corresponding ones of said predicate keys" ("constructing multi-dimensional vectors using said integer representation" is taught by Maarek on Page 5, Paragraph 4 as *For example, if the indexing units are single words, then each word represents an axis in a high-dimensional vector space, where the dimension is equal to the number of words in the collection.*" and Page 7 Last Paragraph and Page

First Paragraph, i.e. *Instead of the typical use of single words as indexing units, our indexing unit consists of a pair of words that are linked by a lexical affinity (LA)).* Note that, in the method of Call, data elements/concepts are identified by integers and data elements/concepts of similar attribute are organized into sets. Thus, it is inherent in Call's method that conceptual nearness is obtained by subtracting corresponding one of predicate keys.

As per claim 96, Maarek teaches the limitation:

"the method further comprising the further step of normalizing said N-dimensional vector representations" (Maarek et al., Page 5 Third Paragraph, i.e.
after normalization.)

Claims 97-98 are rejected on the same basis as claim 95-96 respectively.

12. Claim 94 is rejected under 35 U.S.C. 103(a) as being unpatentable over Maarek in view of Call and further in view of Liddy and further in view of Dorocak and further in view of Anderson et al., (hereinafter "Anderson") (U.S. Patent Application Publication Number 2004/0010501).

As per claim 94, Maarek in view of Call and further Liddy and further in view of Dorocak does not explicitly teach the limitation: "wherein any predicate structure of said M predicate structures that includes more than q arguments omits remaining arguments after q argument positions are filled".

Art Unit: 2162

Anderson teaches the limitation:

"wherein any predicate structure of said M predicate structures that includes more than q arguments omits remaining arguments after q argument positions are filled" (Anderson, Paragraph 0039, i.e., *The lookup/interpolation module 13 may truncate the inquiry parameter values to be within the range of the system parameter values*).

At the time the invention was made, it would have been obvious to a person of ordinary skill in the art to add the feature of truncating parameters as taught by Anderson to the method of Maarek in view of Call, further in view of Liddy and further in view of Dorocak so that the resultant method would comprise parameter/argument truncation. One would have been motivated to do so because it is a notoriously well-known practice in the art of programming languages that excess arguments are truncated (omitted) when number of arguments to a routine is more than the predefined number of arguments of the routine.

Conclusion

13. Applicant's arguments have been considered but are not persuasive.

Accordingly, **THIS ACTION IS MADE FINAL**. See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the

Art Unit: 2162

shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

Contact Information

14. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Dennis Myint whose telephone number is (571) 272-5629. The examiner can normally be reached on 8:30 AM - 5:30 PM Monday-Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene can be reached on (571) 272-4107. The fax phone number for the organization where this application or proceeding is assigned is 571-273-5629. Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

hj
Dennis Myint
Examiner
AU-2162

John E. Breene
JOHN BREENE
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 2100